Compositional data analysis of household food waste in Denmark

Edjabou, Vincent Maklawe Essonanawe; Pivnenko, Kostyantyn; Petersen, Claus; Scheutz, Charlotte; Astrup, Thomas Fruergaard

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Compositional data analysis of household food waste in Denmark

M. E. Edjabou 1, K. Pivnenko 1, C. Petersen 2, C. Scheutz 1, T. Astrup 1
Department of Environmental Engineering, Technical University of Denmark (DTU), Denmark
[2] EconetAS, Copenhagen, Denmark

Abstract

Food waste is a growing public concern because the food production and distribution exert enormous pressure on natural resources such as land, water and energy, and leads to significant environmental, societal and economic impacts. Thus, the European Commission has aimed to reduce to 50% the total amount of discarded edible food waste by 2020 within the European Union (EU) Member States.
Reliable data on food waste and a better understanding of the food waste generation patterns are crucial for planning the avoidable food waste reduction and an environmental sound treatment of unavoidable food waste. Although, food waste composition carries relative information, no attempt was made to analysis food waste composition as compositional data. Thus the relationship between food waste fractions has been analysed by mean of Pearson correlation test and log-ratio analysis. The food waste data was collected by sampling and sorting residual household waste in Denmark. The food waste was subdivided into three fractions: (1) avoidable vegetable food waste, (2) avoidable animal-derive food waste, and (3) avoidable food waste. The correlation was carried out using: (a) the amount of food waste (kg per household per week), (b) percentage composition of food waste based on the total food waste, and (c) percentage composition of food waste based on the total residual household waste. The Pearson correlation test showed different results when different datasets are used, whereas the log-ratio analysis showed the same results for all the three datasets.

Kew words: Compositional data analysis, food waste, Residual household waste, log-ratio, Pearson correlation.

1 Introduction

Food waste is a growing public concern because the food waste production and distribution exert enormous pressure on scarce natural resources such as land, water and energy, and leads to serious environmental, social and economic impacts. Thus, the European Commission has aimed to reduce to 50% the total amount of discarded edible food waste by 2020 within the European Union (EU) Member States.
Reliable data on the amount of food waste is essential to map the current food waste situation and to assess the performance against the food waste reduction targets. On the other, data on food waste composition is crucial to analyse the proportion and type of food products which are wasted, in order to propose adapted and practical solutions of reducing household food waste such as public awareness program, food leftovers recipes, etc.
Food waste composition data carry relative information and they are closed datasets because percentages of individual food waste fractions are positive and sum up to 100 (Reimann et al. 2008; Buccianti and Pawlowsky-Glahn 2011). For this reason, they are considered as compositional data. However, although numerous studies have investigated the composition of household food waste (Lebersorger and Schneider 2011; WRAP 2009), none has analysed household food waste composition as compositional data.
The objective of this paper is identify general patterns of Danish household food waste generation by mean of correlation test between food waste fractions.

2 Materials and Methods

2.1 Sampling of food waste from the Danish households

In order to determine the amount and composition of the Danish household, residual household waste from 800 households located in different municipalities was sampled following the method described by Nordtest (1995). Here, the residual household waste is the remaining mixed waste after source sorting of recyclables materials such as Waste Electrical and Electronic Equipment (WEEE), hazardous waste, gardening waste and bulky waste (Christensen and Matsufuji 2010).. In Denmark, residual household waste is incinerated to generate heat and electricity (Danish Government 2013). The residual household waste was collected following the normal waste collection schedule to avoid any change in the behaviour of household related to their food waste generation. The residual household waste was generated during one week. The residual waste was weighted and sorted according to the methodology described by Edjabou et al (2015) and Lagerkvist et al. (2011).

2.2 Food waste fractions

Food waste is subdivided into avoidable food waste and unavoidable food waste as shown in Table 1.
Avoidable food waste refers to edible food that could be eaten but was disposed of after being rotted or not. Avoidable food waste is subdivided into avoidable vegetable food waste (VeAvoid) and avoidable animal derived food waste (AnAvoid). Avoidable vegetable food waste refers to vegetable products such as fruits, vegetables, and cereals. Avoidable animal-derived food waste refers to meat products, fish products, and dairy products and mixed of animal and vegetable products for example pizza, salad.
Unavoidable food waste fractions refer to food scraps or residues which could not be eaten. This includes for example, bones, carcasses, and certain organs in meat products that are not commonly eaten, egg shells, peels, fruit skin (pineapple, banana…), apple cores, coffee ground… As a result, we have three food waste fractions: (1) avoidable vegetable food waste, (2) avoidable animal-derive food waste, and (3) avoidable food waste

Table 1: Characterisation of food waste

Main fractions	Types	Waste components
Avoidable	Vegetable	Rice, pasta, potatoes, …
		Fresh fruit, fresh carrots and potatoes, bread, cereals,…
	Animal-derived	Cooked eggs, rest of food containing meat, fish,…
		Eggs not cooked, dairy products, not cooked meat and fish, …
Unavoidable		Leftovers containing meat, fish, skins and bones,…
		Residues from fruits, vegetables, coffee grounds

Source: (Edjabou et al. 2013)

3 A compositional Data sets

In the present study, VeAvoid is avoidable vegetable food waste, AnAvoid is avoidable animal-derived food waste and UAvoid is unavoidable food waste.
Figure 1 shows the distribution of the amount of the food waste fractions generated by Danish households presented as kg wet mass food waste per household per week (in the left), percentage composition of individual food waste based on total food waste (middle of Figure 1) and percentage composition of individual food waste fractions based on the total residual household waste (right of Figure 1). The composition is presented as percentage of wet waste.
In average, 131±116 kg wet mass of avoidable food waste per household per year of which avoidable vegetable food waste accounted for 92±86 kg and avoidable animal-derived food waste accounted for 40±47 kg.
The waste composition showed that residual household waste consisted of avoidable vegetable food waste (15±11%), avoidable animal–derived food waste (7±6%), and unavoidable food waste (19±12%) (Figure 1, in the middle). Moreover, household food waste consisted of avoidable vegetable food waste (37±20%), avoidable animal-derived food waste (17±15%) and unavoidable food waste (47±14%) (Figure 1 in the right)

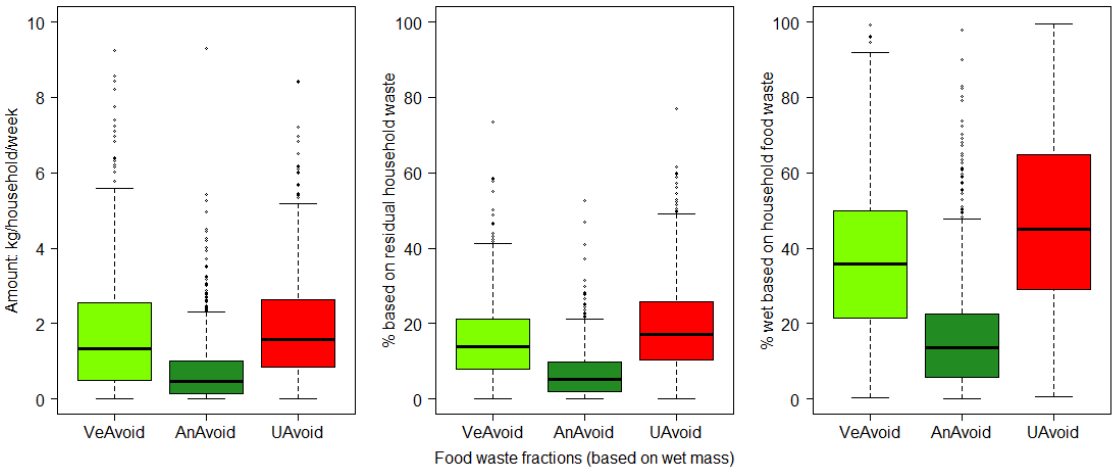

Figure 1: Distribution of the Danish household food waste

As a result, the food waste data are presented as three different datasets. This leads to the following questions:

(1) Which dataset is suitable to detect the true correlation between food waste fractions?
(2) Is the total amount household food waste relevant with respect to the relationship between food waste fractions?

These questions correspond to the principles of compositional analysis (Buccianti and Pawlowsky-Glahn 2011; Aitchison 1994; Filzmoser and Hron 2008; Egozcue and Pawlowsky-Glahn 2011)

4 Correlation between food waste fractions

4.1 Pearson correlation between food waste fractions

Table 2: Correlation coefficients between food waste fractions

	VeAvoid	AnAvoid	UAvoid	
AnAvoid	.	0.46	0.17	Amount[1]
	-	-0.14	-0.53	FW[2]
	-	0.09	-0.23	RHW[3]
VeAvoid	-	-	0.25	Amount
	-	-	-0.76	FW
	-	-	-0.22	RHW

[1]Amount of food waste (kg per household per week);
[2]Composition data of food waste based on the total household food waste
[3]Composition data of food waste based on the total residual household waste

Table 2 shows the correlation coefficients between food waste fractions based on (a) amount (Amount), (b) the composition of food waste (FW) and (c) the composition of the total residual household waste. The correlation coefficients differ considerable from datasets and suggest contradictory conclusions. For example, the correlation coefficient between avoidable animal-derived food waste and unavoidable food waste was 0.17 for the datasets using the amount (kg per household per week) of food waste. This correlation coefficient was -0.53 for food waste composition dataset and -0.23 for total residual household waste composition dataset. The negative correlation coefficients was due to the unit sum constraint(van den Boogaart and Tolosana-Delgado 2013; Egozcue and Pawlowsky-Glahn 2011). Moreover, the correlation based on the composition of food waste and residual household waste may be misleading, since it is based on spurious correlation (Filzmoser and Hron 2008).

4.2 Log-ratio analysis of food waste fractions

Table 3: % total variance and variation array. Upper triangle shows pairwise log-ratio variances in percentage of total variance; lower triangle shows pairwise log-ratio means

	AnAvoid	VeAvoid	UAvoid	% clr variances	
AnAvoid	-	2.95	3.73	1.25	Amount
	-	2.95	3.73	1.25	FW
	-	2.95	3.73	1.25	RHW
VeAvoid	1.18	-	2.14	0.72	Amount
	1.18	-	2.14	0.72	FW
	1.18	-	2.14	0.72	RHW
Uavoid	1.53	0.35		0.98	Amount
	1.53	0.35		0.98	FW
	1.53	0.35		0.98	RHW
Means			Total variance	2.94	Amount
			-	2.94	FW
			-	2.94	RHW

The log-ratio variance matrix of each pair of variables was computed and shown in Table 3. This variance matrix was computed using the following equations (Egozcue and Pawlowsky-Glahn 2011):

$$T = \begin{pmatrix} t_{11} t_{12} \cdots t_{1D} \\ t_{21} t_{22} \ldots t_{2D} \\ \vdots\vdots \ddots \\ t_{D1} t_{D2} \ldots . t_{DD} \end{pmatrix}, t_{ij} = \mathrm{var}\left(\ln \frac{xi}{xj} \right) \quad \text{(Eq. 1)}$$

This variance matrix describes consistently the dispersion in these three datasets and shows the pairwise log-ratio means and the percentage of total variance represented by pairwise log-ratio variances of the food waste fractions.

The variance of log-ratio is highly correlated when the variance of log-ratio is zero (Egozcue and Pawlowsky-Glahn 2011). The permutation test was applied to test the significance of the variance of log ratio between pairs. The p-value of the permutation test showed a significant correlation between these three food waste fractions. This results suggest that there is a relationship between avoidable food waste and unavoidable food waste.

Conclusion

Log-ratio analysis is a powerful tool to assess consistently the correlation and variation between food waste fractions. It enables to discover the true relationship between food waste fractions regardless the dataset used. Log-ratio analysis provides the same result of the three datasets: (a) the amount of food waste (kg per household per week), (b) percentage composition of food waste based on the total food waste, and (c) percentage composition of food waste based on the total residual household waste. However, the Pearson correlation test provides different results of the three food waste datasets.

Acknowledgements and appendices

The authors wish to acknowledge the Danish Strategic Research Council for financing this study via the IRMAR (Integrated Resource Management & Recovery) project (nr. 11-116775).

References

Aitchison, John. 1994. “A Concise Guide to Compositional Data Analysis.” *Lecture Notes-Monograph Series* 24: 73–81. http://www.jstor.org/stable/4355794.

Buccianti, Antonella, and Vera Pawlowsky-Glahn. 2011. *Compositional Data Analysis*. Edited by Vera Pawlowsky-Glahn and Antonella Buccianti. *Compositional Data Analysis: Theory and Applications*. Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9781119976462. http://doi.wiley.com/10.1002/9781119976462.

Christensen, TH, and Y Matsufuji. 2010. “Source Segregation and Collection of Source□Segregated Waste.” *Solid Waste Technology & …*. http://onlinelibrary.wiley.com/doi/10.1002/9780470666883.ch22/summary.

Danish Government. 2013. *Denmark without Waste: Recycle More -Incinerate Less*. Copenhagen, Denmark: Danish Ministry of the Environment. doi:978-87-03026-59-5. http://mim.dk/media/mim/67848/Ressourcestrategi_UK_web.pdf.

Edjabou, Maklawe Essonanawe;, Morten Bang; Jensen, Ramona; Götze, Kostyantyn; Pivnenko, Claus; Petersen, Charlotte; Scheutz, and Thomas Fruergaard Astrup. 2015. “Municipal Solid Waste Composition: Sampling Methodology , Statistical Analyses , and Case Study Evaluation.” *Waste Management* 36: 12–23. doi:10.1016/j.wasman.2014.11.009. http://dx.doi.org/10.1016/j.wasman.2014.11.009.

Edjabou, Vincent Maklawe Essonanawe, C. Petersen, Charlotte Scheutz, and Thomas Fruergaard Astrup. 2013. “Characterization of Household Food Waste in Denmark.” In .

Egozcue, J, and V Pawlowsky□Glahn. 2011. “Basic Concepts and Procedures.” Edited by Vera Pawlowsky-Glahn and Antonella Buccianti. *Compositional Data Analysis* (September 23): 12–28. doi:10.1002/9781119976462. http://doi.wiley.com/10.1002/9781119976462.

Filzmoser, Peter, and Karel Hron. 2008. “Correlation Analysis for Compositional Data.” *Mathematical Geosciences* 41 (8) (October 15): 905–919. doi:10.1007/s11004-008-9196-y. http://link.springer.com/10.1007/s11004-008-9196-y.

Lagerkvist, Anders, Holger Ecke, and Thomas H Christensen. 2011. “Waste Generation and Characterization Waste Characterization : Approaches and Methods.” In *Solid Waste Technology & Management, Volume 1 & 2*, edited by T.H. Christensen. Chichester, UK. doi:10.1002/9780470666883.ch5.

Lebersorger, S, and F Schneider. 2011. “Discussion on the Methodology for Determining Food Waste in Household Waste Composition Studies.” *Waste Management (New York, N.Y.)* 31 (9-10): 1924–33. doi:10.1016/j.wasman.2011.05.023. http://www.ncbi.nlm.nih.gov/pubmed/21705207.

Nordtest. 1995. “Municipal Solid Waste: Sampling and Characterisation.” NT ENVIR 001. Nordtest Method. Espoo, Finland. http://www.nordtest.info/images/documents/nt-methods/environment/NT envir 001_Solid waste, municipal_Sampling and characterisation_Nordtest Method.pdf.

Reimann, Clemens, Peter Filzmoser, Robert G. Garrett, and Rudolf Dutter. 2008. *Statistical Data Analysis Explained*. *Statistical Data Analysis Explained: Applied Environmental Statistics With R*. Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9780470987605. http://doi.wiley.com/10.1002/9780470987605.

Van den Boogaart, K. Gerald, and Raimon Tolosana-Delgado. 2013. *Analyzing Compositional Data with R*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-36809-7. http://link.springer.com/10.1007/978-3-642-36809-7.

WRAP. 2009. *Household Food and Drink Waste in the UK*. *October*. Banbury, UK. doi:10.1111/j.1467-3010.2011.01924.x.